

C.3



NUMERICAL INTEGRATION OF FIRST-ORDER STIFF DIFFERENTIAL EQUATIONS

F. C. Loper and W. J. Phares

ARO, Inc.

February 1966

PROPERTY OF U. S. AIR FORCE
AEDC LIBRARY
AF 40(600)1200

Distribution of this document is unlimited.

**ENGINEERING SUPPORT FACILITY
ARNOLD ENGINEERING DEVELOPMENT CENTER
AIR FORCE SYSTEMS COMMAND
ARNOLD AIR FORCE STATION, TENNESSEE**

PROPERTY OF U. S. AIR FORCE
AEDC LIBRARY
AF 40(600)1200

NOTICES

When U. S. Government drawings specifications, or other data are used for any purpose other than a definitely related Government procurement operation, the Government thereby incurs no responsibility nor any obligation whatsoever, and the fact that the Government may have formulated, furnished, or in any way supplied the said drawings, specifications, or other data, is not to be regarded by implication or otherwise, or in any manner licensing the holder or any other person or corporation, or conveying any rights or permission to manufacture, use, or sell any patented invention that may in any way be related thereto.

Qualified users may obtain copies of this report from the Defense Documentation Center.

References to named commercial products in this report are not to be considered in any sense as an endorsement of the product by the United States Air Force or the Government.

NUMERICAL INTEGRATION OF FIRST-ORDER
STIFF DIFFERENTIAL EQUATIONS

F. C. Loper and W. J. Phares
ARO, Inc.

Distribution of this document is unlimited.

FOREWORD

The work reported herein was sponsored by the Arnold Engineering Development Center (AEDC), Air Force Systems Command (AFSC), under Program Element 62405214, Project 6951, Task 695102.

The results of the research presented were obtained by ARO, Inc. (a subsidiary of Sverdrup and Parcel, Inc.), contract operator of the AEDC, AFSC, Arnold Air Force Station, Tennessee, under Contract AF40(600)-1200. The research was performed from January until June, 1965, under ARO Project No. RW2408, and the manuscript was submitted for publication on November 17, 1965.

The solution of the mathematical problem to which this report is addressed was stimulated by a requirement to solve a set of ordinary, first-order, nonlinear, stiff differential equations which describe chemical reaction rates applicable to hydrogen-air reactions. The authors wish to thank R. P. Rhodes, Jr., who proposed the problem which led to the development of this technique, and who worked jointly with the authors in solving the proposed problem.

This technical report has been reviewed and is approved.

Marion L. Laster
Propulsion Division
DCS/Research

Donald R. Eastman, Jr.
DCS/Research

ABSTRACT

A method is presented for numerically integrating a system of stiff, first-order differential equations. This method is based on transforming the set of dependent variables so that the resulting system will not be stiff; the transformed system is then integrated by the Runge-Kutta method. The resulting procedure is often appreciably faster than classical methods in that a much larger step size is allowable with nominal increase in step computation time. Applications and results are discussed for systems of various order, including a system of six chemical rate equations.

CONTENTS

	<u>Page</u>
ABSTRACT.	iii
I. INTRODUCTION	1
II. DEFINITIONS AND NOMENCLATURE.	1
III. DERIVATION OF THE TRANSFORMATION	3
IV. APPLICATION TECHNIQUES	6
V. DISCUSSION OF RESULTS.	8
VI. CONCLUSIONS	9
REFERENCES	9

ILLUSTRATIONS

Figure

1.	Errors for a Single Equation.	11
2.	Errors for a System of Two Equations	12

SECTION I INTRODUCTION

Although systems of differential equations having the property of being stiff have been used as mathematical models of certain physical phenomena almost since the invention of the calculus, only in recent years have such systems been categorized according to this property and been formally recognized as a class of equations, the numerical solution of which often eludes the computer when sought by classical means. Since 1951, various works have been published on the subject (Refs. 1 through 5).

The primary reason that equations in this class present such difficulty is that an exceedingly small integration step size is sometimes required, making classical integration techniques impractical even on the most sophisticated computing machines.

The object of this presentation is to formulate a method that can be used efficiently on a high speed digital computer to obtain the solution of a system of first-order, ordinary, stiff differential equations. The method described is straightforward and practical when applied to stiff equations resulting from many physical problems.

The above objective may be realized by transforming the dependent variables in such a manner that the resulting system of equations will not be stiff in a neighborhood of the value of the independent variable at which the transformation was made. The Runge-Kutta method will be used to integrate the resulting system.

SECTION II DEFINITIONS AND NOMENCLATURE

When presenting a similar technique, some authors develop their theory exclusively for a single equation with one dependent variable, and later state that the extension of their method to a system of such equations is "obvious". It is the contention of the authors of this presentation that any such extension is almost always ambiguous. It is also believed that with little or no sacrifice to clarity, the method can be developed in vector notation, leaving no doubt concerning how the method should be applied to the general case. Following these convictions, the nomenclature introduced below will be used throughout this document:

1. Roman characters will be used to denote real scalar quantities.
2. Greek characters will denote numbers which are, in general, complex scalars.
3. An (\rightarrow) will indicate an $n \times 1$ column vector, and $\vec{\phi}$ denotes the null vector.
4. An $(-)$ will denote an $n \times n$ matrix; $\vec{\phi}$ and \vec{I} denote the null matrix and identity matrix, respectively.
5. The notation $\text{Re } (\lambda)$ will mean the real part of the complex number, λ .

Now consider the system of equations

$$\frac{dy_i}{dx} = F_i(x, y_1, y_2, \dots, y_n); y_i(x_0) = y_{i0} \quad (1)$$

for $i = 1, \dots, n$. More conveniently, Eq. (1) can be written

$$\frac{d\vec{y}}{dx} = \vec{F}(x, \vec{y}), \quad \vec{y}(x_0) = \vec{y}_0 \quad (2)$$

where

$$\vec{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix} \quad \vec{F}(x, \vec{y}) = \begin{bmatrix} F_1(x, \vec{y}) \\ F_2(x, \vec{y}) \\ \vdots \\ F_n(x, \vec{y}) \end{bmatrix}$$

In order to get Eq. (2) in an applicable form, define the matrix

$$-\vec{f}(x, \vec{y}) = \begin{bmatrix} \frac{\partial F_1}{\partial y_1}(x, \vec{y}) & \frac{\partial F_1}{\partial y_2}(x, \vec{y}) & \dots & \frac{\partial F_1}{\partial y_n}(x, \vec{y}) \\ \frac{\partial F_2}{\partial y_1}(x, \vec{y}) & \frac{\partial F_2}{\partial y_2}(x, \vec{y}) & \dots & \frac{\partial F_2}{\partial y_n}(x, \vec{y}) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial F_n}{\partial y_1}(x, \vec{y}) & \frac{\partial F_n}{\partial y_2}(x, \vec{y}) & \dots & \frac{\partial F_n}{\partial y_n}(x, \vec{y}) \end{bmatrix} \quad (3)$$

It will be assumed throughout that $\tilde{f}(x, \vec{y})$ is nonsingular, this method being inapplicable at any point on the solution of Eq. (2) for which this is not true. Equation (2) can now be written

$$\frac{d\vec{y}}{dx} + \tilde{f}(x, \vec{y})\vec{y} = \vec{g}(x, \vec{y}) \quad , \quad \vec{y}(x_0) = \vec{y}_0 \quad (4)$$

where

$$\vec{g}(x, \vec{y}) = \vec{F}(x, \vec{y}) + \tilde{f}(x, \vec{y})\vec{y} \quad (5)$$

The form of Eq. (4) is now such that the definition of a stiff differential equation can be given. Definition: Let $x = x_1$ be a value of the independent variable in the region of interest, and let $\lambda_1, \lambda_2, \dots, \lambda_n$ be the eigenvalues of $\tilde{f}(x_1, \vec{y}(x_1))$. If

$$\text{Re}(\lambda_i) > 0$$

for any i , then Eq. (4) is said to be stiff at $x = x_1$.

SECTION III DERIVATION OF THE TRANSFORMATION

It was mentioned previously that the dependent variable in Eq. (4), $\vec{y}(x_0)$, will be transformed so that if $|x - x_0| \leq h$, then the transformed equation will not be stiff. For the remainder of this paper, h is understood to be the integration step size for the Runge-Kutta method (fourth-order) on the transformed equations. The maximum allowable value of h will, of course, depend on the exact nature of the original differential equation; the transformation has allowed the value of h to be significantly increased in all test cases tried so far.

It is pointed out again that the validity of the method depends on the nonsingularity of $\tilde{f}(x, \vec{y})$.

It is convenient to define at this time the matrix

$$\tilde{u}(x) = \tilde{I} + \tilde{f}_0(x - x_0) + \frac{\tilde{f}_0^2(x - x_0)^2}{2!} + \frac{\tilde{f}_0^3(x - x_0)^3}{3!} + \dots \quad (6)$$

which is sometimes referred to as

$$\tilde{u}(x) = \exp(-\tilde{f}_0(x - x_0))$$

for obvious reasons.

The following properties of the infinite matrix series, Eq. (6), can be readily verified independent of \tilde{f}_0 and $(x - x_0)$;

1. The series converges
2. $\tilde{u}(x)$ is nonsingular
3. $\tilde{u}(x) \tilde{f}_0 = \tilde{f}_0 \tilde{u}(x)$

Note also that

$$\frac{d \tilde{u}(x)}{dx} + \tilde{f}_0 \tilde{u}(x) = \tilde{\phi}, \quad \tilde{u}(x_0) = \tilde{I} \quad (7)$$

Now, let $\vec{z}(x)$ be a solution of

$$\frac{d \vec{z}}{dx} + \tilde{f}_0 \vec{z} = \vec{g}_0, \quad \vec{z}_0 = \vec{y}_0 \quad (8)$$

Hence

$$z = \tilde{u}(x) (\vec{y}_0 - \tilde{f}_0^{-1} \vec{g}_0) + \tilde{f}_0^{-1} \vec{g}_0 \quad (9)$$

Defining

$\vec{w}(x) = \vec{y}(x) - \vec{z}(x)$ one gets, by subtracting Eq. (8) from Eq. (4)

$$\frac{d \vec{w}(x)}{dx} + \tilde{f}(x, \vec{y}) \vec{w}(x) = \vec{v}(x, \vec{y}); \quad \vec{w}_0 = \vec{\phi} \quad (10)$$

where

$$\vec{v}(x, \vec{y}) = \vec{g}(x, \vec{y}) - \vec{g}_0 + [\tilde{f}_0 - \tilde{f}(x, \vec{y})] \vec{z}(x) \quad (11)$$

Premultiplying Eq. (10) by $\tilde{u}(x)$ and postmultiplying Eq. (7) by $\vec{w}(x)$ gives

$$\tilde{u}(x) \frac{d \vec{w}}{dx} + \tilde{u} \tilde{f}(x, \vec{y}) \vec{w} = \tilde{u} \vec{v}$$

and

$$\frac{d \tilde{u}}{dx} \vec{w} + \tilde{f}_0 \tilde{u}(x) \vec{w} = \vec{\phi}$$

Subtracting,

$$\begin{aligned} \tilde{u} \frac{d \vec{w}}{dx} - \frac{d \tilde{u}}{dx} \vec{w} + \tilde{u} \tilde{f}(x, \vec{y}) \vec{w} - \tilde{f}_0 \tilde{u}(x) \vec{w} &= \tilde{u} \vec{v} \\ (\tilde{u}^{-1})^2 \left(\tilde{u} \frac{d \vec{w}}{dx} - \frac{d \tilde{u}}{dx} \vec{w} \right) + (\tilde{u}^{-1})^2 \left[\tilde{u} \tilde{f}(x, \vec{y}) \vec{w} - \tilde{f}_0 \tilde{u}(x) \vec{w} \right] &= (\tilde{u}^{-1})^2 \tilde{u} \vec{v} \\ \frac{d}{dx} (\tilde{u}^{-1} \vec{w}) + \tilde{u}^{-1} \left[\tilde{f}(x, \vec{y}) - \tilde{u}^{-1} \tilde{f}_0 \tilde{u} \right] \vec{w} &= \tilde{u}^{-1} \vec{v} \end{aligned} \quad (12)$$

Defining

$$\begin{aligned}\vec{t} &= \vec{u}^{-1} \vec{w} \\ \vec{m} &= \vec{u}^{-1} \left[\vec{f}(x, \vec{y}) - \vec{f}_t \right] \vec{u}\end{aligned}$$

Eq. (12) becomes

$$\frac{d\vec{t}}{dx} + \vec{m} \vec{t} = \vec{u}^{-1} \vec{v}; \quad \vec{t}(x_0) = \vec{\phi} \quad (13)$$

Equation (13) is the desired transformed equation, the relation between \vec{t} and \vec{y} being

$$\vec{y} = \vec{u}(x) \vec{t}(x) + \vec{f}_0^{-1} \left[\vec{I} - \vec{u}(x) \right] \vec{F}_0 + \vec{y}_0 \quad (14)$$

Setting

$$\vec{G}(x, \vec{t}) = \vec{u}^{-1} \vec{v} - \vec{m} \vec{t} = \vec{u}^{-1} \left[\vec{F}(x, \vec{y}) - \vec{F}_0 + \vec{f}_0 (\vec{y} - \vec{y}_0) \right] \quad (15)$$

Eq. (13) can be written as

$$\frac{d\vec{t}}{dx} = \vec{G}(x, \vec{t}); \quad \vec{t}(x_0) = \vec{\phi} \quad (16)$$

An interesting consequence of the transformation is observed if Eq. (2) is linear with constant coefficients. In that case

$$\begin{aligned}\vec{f}(x, \vec{y}) &\equiv \vec{f}_0 \\ \vec{g}(x, \vec{y}) &\equiv \vec{g}_0 \\ \frac{d\vec{t}}{dx} &\equiv \vec{\phi} \\ \vec{t}(x) &\equiv \vec{\phi} \\ \vec{y}(x) &= \vec{f}_0^{-1} (\vec{I} - \vec{u}) \vec{F}_0 + y_0\end{aligned} \quad (17)$$

Notice that Eq. (17) gives the exact solution of the linear equation with constant coefficients.

The fundamental question to be explored at this point, the transformation having been obtained, is whether, and to what extent, Eq. (16) is stiff. Basic to this question is the matrix

$$\left(\frac{\partial G_i}{\partial t_j} \right)$$

which is associated with Eq. (16) in the same manner that \tilde{f} is associated with Eq. (2). On the assumption that all the necessary partial derivatives exist, it can be shown after some manipulation that

$$\left(\frac{\partial G_i}{\partial t_j} \right) = - \tilde{m}$$

As

$$\tilde{m} = \tilde{u}^{-1} \left[\tilde{f}(x, \vec{y}) - \tilde{f}_0 \right] \tilde{u}$$

it follows that the eigenvalues of m are the same as the eigenvalues of

$$\tilde{f}(x, \vec{y}) - \tilde{f}_0$$

Thus, the spectral radius of \tilde{m} can be made as small as desired simply by choosing an x sufficiently close to x_0 . It is assumed here, of course, that $\tilde{f}(x, \vec{y})$ is continuous at $x = x_0$. According to the definition of a stiff equation given earlier, it follows that Eq. (16) is not a stiff equation at $x = x_0 + h$ provided the value of h is not too large. Any failure in the Runge-Kutta integration of Eq. (16) will, therefore, be the result of some other undesirable phenomenon.

SECTION IV APPLICATION TECHNIQUES

One method of applying the transformation to obtain the solution of Eq. (2) would be to solve the transformed Eq. (16) for $\vec{t}(x)$ by the Runge-Kutta method, compute the corresponding $\vec{y}(x)$, update the transformation, and continue in the same fashion on a new interval. While $\vec{t}(x)$ is piecewise discontinuous, this of course has no bearing on the solution of interest, $\vec{y}(x)$.

Having served its purpose, the transformation can now, in a computational sense, be entirely removed from the procedure. Thus, the technique can be thought of in a completely different way, that of solving the untransformed Eq. (2) by a method different from that of Runge-Kutta.

While the above two interpretations of the method would theoretically yield identical answers, the second version seems to be advantageous to the first in practice. This was experienced in actual test runs in which the two interpretations were compared on the basis of run time, accuracy, and simplicity of program logic.

As a result of Eqs. (14) and (16), and the classical fourth-order Runge-Kutta equations (Ref. 6, p. 122), the equations necessary for integrating Eq. (2) with the transformation removed are

for $j = 0$

$$\begin{aligned}x &= x_0 \\ \vec{y} &= \vec{y}_0 \\ \vec{k}_0 &= \vec{\phi}\end{aligned}\tag{18a}$$

for $j = 1$

$$\begin{aligned}x &= x_0 + \frac{h}{2} \\ \vec{y} &= \tilde{f}_0^{-1} \left[\tilde{I} - \tilde{u}(x) \right] \vec{F}_0 + \vec{y}_0 \\ \vec{k}_1 &= \tilde{u}^{-1}(x) \left[\vec{F}(x, \vec{y}) - \vec{F}_0 + \tilde{f}_0 (\vec{y} - \vec{y}_0) \right]\end{aligned}\tag{18b}$$

for $j = 2$

$$\begin{aligned}x &= x_0 + \frac{h}{2} \\ \vec{y} &= \frac{h}{2} \tilde{u}(x) \vec{k}_1 + \tilde{f}_0^{-1} \left[\tilde{I} - \tilde{u}(x) \right] \vec{F}_0 + \vec{y}_0 \\ \vec{k}_2 &= \tilde{u}^{-1}(x) \left[\vec{F}(x, \vec{y}) - \vec{F}_0 + \tilde{f}_0 (\vec{y} - \vec{y}_0) \right]\end{aligned}\tag{18c}$$

for $j = 3$

$$\begin{aligned}x &= x_0 + h \\ \vec{y} &= h \tilde{u}(x) \vec{k}_2 + \tilde{f}_0^{-1} \left[\tilde{I} - \tilde{u}(x) \right] \vec{F}_0 + \vec{y}_0 \\ \vec{k}_3 &= \tilde{u}^{-1}(x) \left[\vec{F}(x, \vec{y}) - \vec{F}_0 + \tilde{f}_0 (\vec{y} - \vec{y}_0) \right]\end{aligned}\tag{18d}$$

at

$$\begin{aligned}x &= x_0 + h \\ \vec{y} &= \frac{h}{6} \tilde{u}(x) (2\vec{k}_1 + 2\vec{k}_2 + \vec{k}_3) + \tilde{f}_0^{-1} \left[\tilde{I} - \tilde{u}(x) \right] \vec{F}_0 + \vec{y}_0\end{aligned}$$

Notice that although the algebra is somewhat more involved, there is a close similarity between Eq. (18) and the corresponding Runge-Kutta equations.

The problem that initially motivated the development of this technique is a system of six chemical rate equations (Ref. 7). However, the matrix $\tilde{f}(x, y)$ associated with this problem was found to be ill-conditioned; hence, a modified version of the method was developed.

Equation (18) in scalar form was applied to each of the six equations individually. In other words, the six equations were, for purposes of making the transformation only, assumed to be uncoupled.

In order to determine whether or not to make the transformation on the i th equation, $\frac{dy_i}{dx} = f_i(x, \vec{y})$, the number $-\frac{\partial f_i}{\partial y_i} \geq 0$ was examined to see if the equation was suitable for integration by the Runge-Kutta Method.

If

$$-\frac{\partial f_i}{\partial y_i} \Delta x \geq 0.5$$

then the i th equation was transformed. If the above inequality did not hold, the i th equation was considered suitable for integration without making the transformation (Ref. 8, p. 198, or Ref. 1). Thus the six equations were solved as a coupled system, but the integration method used on the individual equations was dictated by the above criterion. Also, the test was made at every integration step so that at one value of x , for example, Eqs. (1), (2), (5), and (6) were transformed while Eqs. (2) and (4) were integrated by the Runge-Kutta method. At another value of x , an entirely different combination might very well have prevailed.

SECTION V DISCUSSION OF RESULTS

Several stiff systems for which the analytic solution is known were solved both by the Runge-Kutta method and by the method under consideration here. One such equation is

$$y' + ay = x^2; \quad y(0) = 0$$

A comparison of the errors, the absolute value of the difference between the exact answer and the approximation, is shown in Fig. 1 for $a = 100$ and $a = 1000$.

Figure 2 shows a similar graph for the Euclidian norm of the error vector associated with the system.

$$-y_1' + a_{11} y_1 + a_{12} y_2 = b_1 x; \quad y_1(0) = y_1^0$$

$$-y_2' + a_{21} y_1 + a_{22} y_2 = b_2 x^2; \quad y_2(0) = y_2^0$$

No analytic solution was available for analyzing the success of the method as applied to the chemical rate equations. Essentially, the same solution was found using both the Runge-Kutta method proper and the transformed method, a considerably larger Δx being admissible in the latter case. These answers were also compared with solutions from another source (Ref. 7) as further confirmation of their validity.

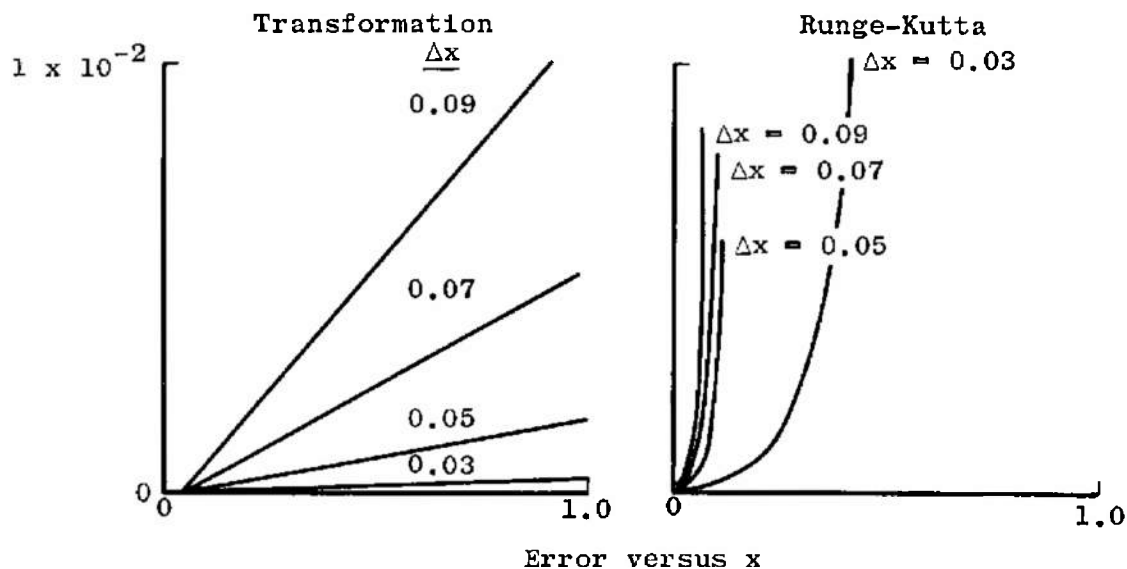
SECTION VI CONCLUSIONS

The method discussed herein for integrating a system of stiff, first-order differential equations has been found to be practical and expedient in all cases on which the method has been tested. Answers comparable to those indicated by the Runge-Kutta method were obtained with a considerably larger integration step size.

REFERENCES

1. Emanuel, G. "Numerical Analysis of Stiff Equations." SSD-TDR-63-380 (AD 431750), January 1964.
2. Emanuel, G. "Problems Underlying the Numerical Integration of the Chemical and Vibrational Rate Equations in a Near-Equilibrium Flow." AEDC-TDR-63-82 (AD 400745), March 1963.
3. Curtiss, C. F. and Hirschfelder, J. O. "Integration of Stiff Equations." National Academy of Sciences, Proceedings. Vol. 38, Page 235, 1952.
4. Moretti, L. "A New Technique for the Numerical Analysis of Non-Equilibrium Flows." GASL TR 412.
5. Treanor, C. E. "A Method for the Numerical Integration of Coupled First Order Differential Equations with Greatly Different Time Constants." CAL AG-1729-A-4, 1964.
6. Henrici, P. "Discrete Variable Methods in Ordinary Differential Equations." John Wiley & Sons, Inc., New York, 1962.
7. Libby, P. A., Pergament, H. S., and Bloom, M. H. "A Theoretical Investigation of Hydrogen-Air Reactions." GASL TR 250, 1961.
8. Hamming, R. W. "Numerical Methods for Scientists and Engineers." McGraw-Hill Book Company, Inc., New York, 1962.

$$y' + 100y = x^2$$



$$y' + 1000y = x^2$$

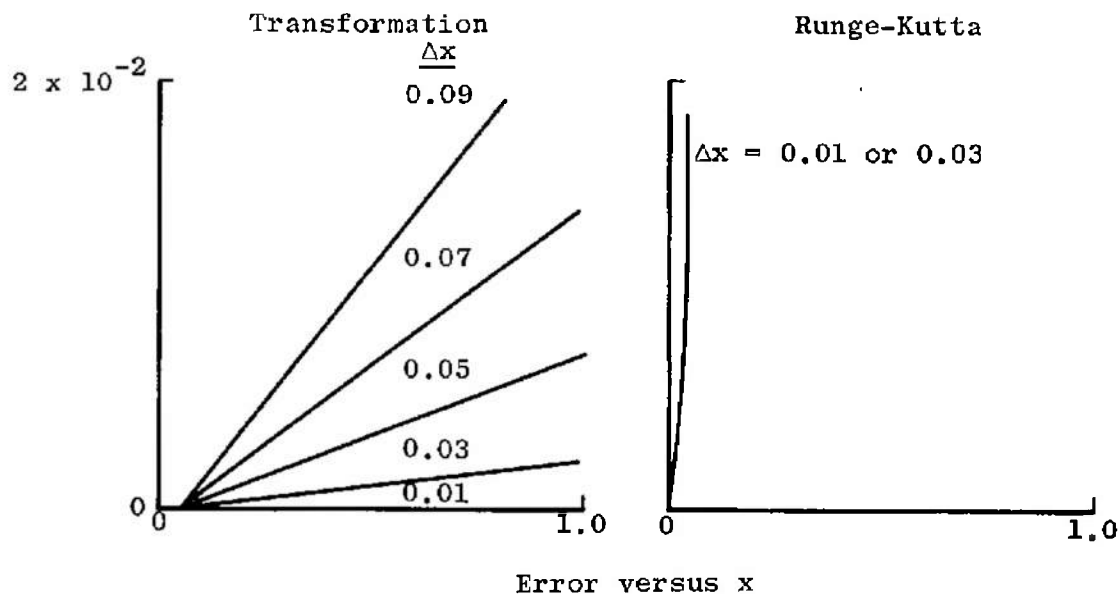


Fig. 1 Errors for a Single Equation

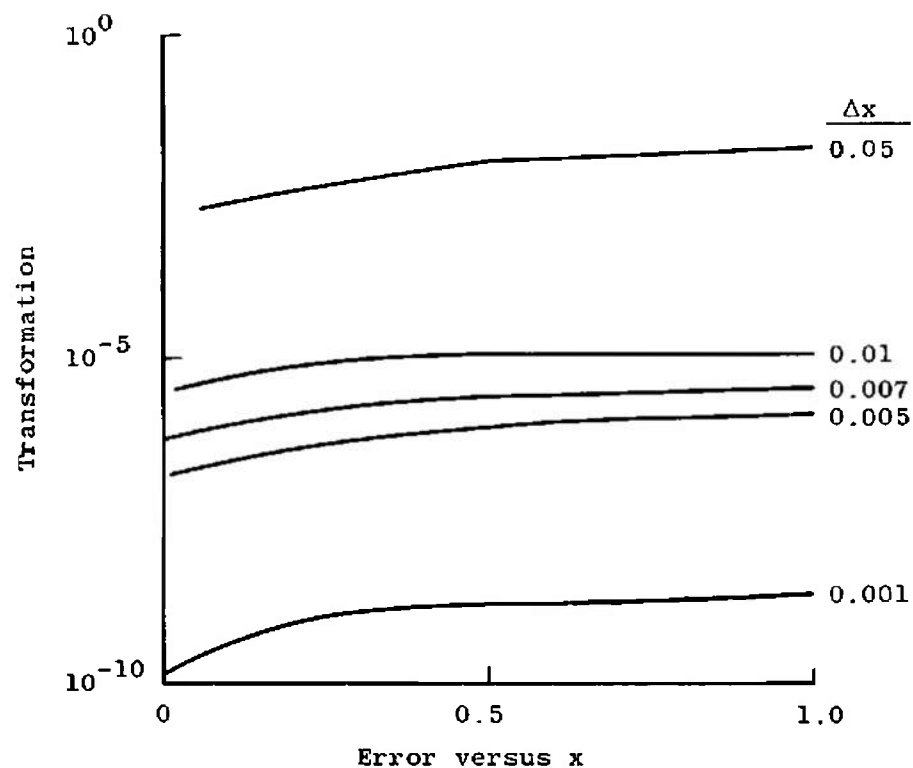
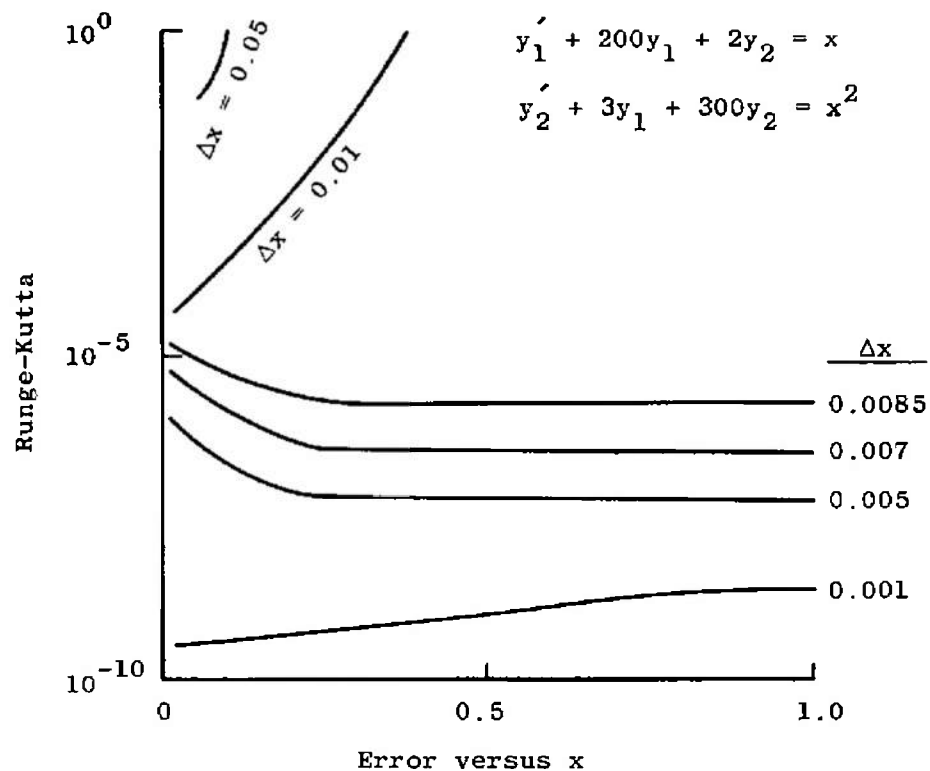


Fig. 2 Errors for a System of Two Equations

UNCLASSIFIED

Security Classification

DOCUMENT CONTROL DATA - R&D

(Security classification of title, body of abstract and indexing annotation must be entered when the overall report is classified)

1 ORIGINATING ACTIVITY (Corporate author) Arnold Engineering Development Center ARO, Inc., Operating Contractor Arnold AF Station, Tennessee		2a REPORT SECURITY CLASSIFICATION UNCLASSIFIED	
		2b GROUP N/A	
3 REPORT TITLE NUMERICAL INTEGRATION OF FIRST-ORDER STIFF DIFFERENTIAL EQUATIONS			
4 DESCRIPTIVE NOTES (Type of report and inclusive dates) N/A			
5 AUTHOR(S) (Last name, first name, initial) Loper, F. C. and Phares, W. J., ARO, Inc.			
6 REPORT DATE February 1966		7a TOTAL NO. OF PAGES 18	7b NO. OF REFS 8
8a CONTRACT OR GRANT NO AF 40(600)-1200		9a ORIGINATOR'S REPORT NUMBER(S) AEDC-TR-65-262	
b. PROJECT NO. 6951		9b. OTHER REPORT NO(S) (Any other numbers that may be assigned this report) N/A	
c Program Element 62405214			
d			
10 AVAILABILITY/LIMITATION NOTICES Qualified users may obtain copies of this report from DDC.			
11 SUPPLEMENTARY NOTES N/A		12 SPONSORING MILITARY ACTIVITY Arnold Engineering Development Center Air Force Systems Command Arnold AF Station, Tennessee	
13 ABSTRACT A method is presented for numerically integrating a system of stiff, first-order differential equations. This method is based on transforming the set of dependent variables so that the resulting system will not be stiff; the transformed system is then integrated by the Runge-Kutta method. The resulting procedure is often appreciably faster than classical methods in that a much larger step size is allowable with nominal increase in step computation time. Applications and results are discussed for systems of various order, including a system of six chemical rate equations. (U)			

KEY WORDS

LINK A

LINK

LINK C

ROLE

WT

ROLE

NT

ROLE

WT

Security Classification